

PAPER • OPEN ACCESS

Pseudo quantum advantages in perceptron storage capacity

To cite this article: Fabio Benatti *et al* 2026 *J. Phys. A: Math. Theor.* **59** 145203

View the [article online](#) for updates and enhancements.

You may also like

- [Eye tracking in AI-assisted AR glasses enabled by Pancharatnam-Berry optical elements](#)
Luke Benoit, Hosna Tajvidi Safa, Yefu Zhang et al.
- [Generation of High Order Harmonics in Quantum Vacuum for 4-Dipole Geometry](#)
Pavel Sasorov and Sergei V Bulanov
- [Emerging trends in non-contact ultrasonic sensing with electromagnetic acoustic transducer \(EMAT\)](#)
Hirotosugu OGI



PAPER

Pseudo quantum advantages in perceptron storage capacity

OPEN ACCESS

RECEIVED

3 November 2025

REVISED

20 February 2026

ACCEPTED FOR PUBLICATION

23 March 2026

PUBLISHED

7 April 2026

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence.

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



Fabio Benatti^{1,2} , Masoud Gharahi^{1,7} , Giovanni Gramegna^{3,4} , Stefano Mancini^{5,6} and Vincenzo Parisi^{5,6,8,*}

¹ Department of Physics, University of Trieste, Strada Costiera 11, I-34151 Trieste, Italy

² Istituto Nazionale di Fisica Nucleare, Sezione di Trieste, Strada Costiera 11, I-34151 Trieste, Italy

³ Dipartimento di Fisica, Università degli Studi di Bari, I-70126 Bari, Italy

⁴ Istituto Nazionale di Fisica Nucleare, Sezione di Bari, I-70126 Bari, Italy

⁵ School of Science and Technology, University of Camerino, Via Madonna delle Carceri, 9, Camerino I-62032, Italy

⁶ Istituto Nazionale di Fisica Nucleare, Sezione di Perugia, via A. Pascoli, I-06123 Perugia, Italy

⁷ Present address: Faculty of Physics, Astronomy and Applied Computer Science, Jagiellonian University, 30–348 Kraków, Poland.

⁸ Present address: CONCEPT Lab, Fondazione Istituto Italiano di Tecnologia, via E Melen 83, 16152 Genova, Italy.

* Author to whom any correspondence should be addressed.

E-mail: vincenzo.parisi@unicam.it, benatti@ts.infn.it, masoud.gharahi@gmail.com, giovanni.gramegna@uniba.it and stefano.mancini@unicam.it

Keywords: quantum perceptron, storage capacity, quantum neural networks

Abstract

We investigate a generalized quantum perceptron architecture characterized by an oscillating activation function with a tunable frequency ranging from zero to infinity. Employing analytical techniques from statistical mechanics, we derive the optimal storage capacity and demonstrate that the classical result is recovered in the limit of vanishing frequency. As the frequency increases, however, the architecture exhibits enhanced quantum storage capabilities. Notably, this improvement stems solely from the specific form of the activation function and, in principle, could be emulated within a classical framework. Accordingly, we refer to this enhancement as a *pseudo quantum advantage*.

1. Introduction and motivation

Recent advancements in quantum computing have enabled the implementation of machine learning concepts on quantum hardware [1]. This development raises the prospect of quantum neural networks outperforming their classical counterparts, offering enhanced storage capacity and superior information processing capabilities. In the classical framework, using statistical mechanical tools, deep connections among neural networks, spin glasses, and information processing have been uncovered [2, 3]. A significant advantage of the statistical mechanics approach is its ability to extract global, macroscopic features of physical systems without requiring detailed knowledge of their microscopic details. This methodology has also enabled the characterization of artificial neural networks without requiring prior knowledge of specific learning rules. Instead, it approaches the problem by treating network weights as random variables, a framework often referred to as *Gardner's program* [4–6].

In the quantum setting, the statistical approach has already been used to estimate the storage capacity of continuous and discrete versions of quantum perceptrons; namely, of the fundamental building blocks of quantum neural networks. Yet, the models that have been considered show apparently opposite results regarding their storage capabilities. Indeed, the results in [7, 8] indicate that quantum advantages in storage properties are unlikely. There, a natural quantum encoding of the classical patterns and a binary classification rule of the measurement outcomes are provided which yield an optimal storage capacity always bounded from above by the maximal classical storage capacity of $\alpha_c = 2$. The reason behind such a behavior is due to the fuzziness injected into the procedure by the non-perfect distinguishability of the quantum states encoding the classical patterns and by the intrinsic randomness of the measurement results upon which the classification of the quantumly encoded classical patterns is operated.

Conversely, the results in [9] suggest that the quantum storage capacity may become twice its classical counterpart, yet this result stems from a sign ambiguity inherent to the measurement process in

the employed perceptron model. Furthermore, in [10], the authors investigated a quantum perceptron implemented on a quantum circuit using a repeat until success method, finding a storage capacity almost five times larger than the classical one. Again, this advantage ought to be ascribed to the highly non-linear form of the activation function resulting from the employed quantum perceptron model.

In this work, taking inspiration from the latter analysis, we apply Gardner's program to a broadly defined quantum perceptron architecture [11] featuring an oscillating activation function whose frequency can range from zero to infinity. We compute analytically the storage capacity and show that, while the classical result is recovered at vanishing frequency, increasing the frequency yields enhanced—and even diverging—quantum storage capacity. However, since this effect arises solely from the form of the activation function, it can, in principle, be replicated within a classical framework. Therefore, we refer to this enhancement in storage capacity as a *pseudo quantum advantage*.

2. Basic tools

This section deals with the main tools and techniques necessary for our later purposes. We start by recalling the main features of a so-called *classical perceptron* [12–16]. Then, we introduce the notion of storage capacity as an appropriate parameter to characterize the perceptron performances. Finally, we outline the essential features of Gardner's approach to evaluate the storage capacity [2–6].

2.1. The classical perceptron

Artificial neurons are the fundamental building blocks of an artificial neural network. From the mathematical point of view, the output of an artificial neuron can be modeled by a map

$$\mathbb{R}^N \ni \mathbf{x} \mapsto f(\mathbf{w} \cdot \mathbf{x} + b) \in \mathbb{R}, \quad (1)$$

where $\mathbf{x} = (x_1, \dots, x_N)$, $\mathbf{w} = (w_1, \dots, w_N)$ represent the *input patterns* and the *vector of weights*, respectively, while $\mathbf{w} \cdot \mathbf{x} = \sum_{i=1}^N w_i x_i$. The constant term $b \in \mathbb{R}$ is the so-called *bias*, while $f: \mathbb{R} \rightarrow \mathbb{R}$ is a non-linear function—i.e. the *activation function*—which determines whether the artificial neuron is active or not. In the case where the activation function is chosen to be the Heaviside theta function—that is, $f(x) = \Theta(x) = 1$ if $x > 0$, and zero otherwise—or the sign function, $\text{sgn}(x)$, the artificial neuron is called *perceptron* [12–16].

Let $f(x) \equiv \text{sgn}(x)$ and, for simplicity, set the bias $b = 0$ as in the standard task addressed by a classical perceptron. The latter is the so-called binary classification problem, consisting in the assignment of a given input vector to one of two possible classes [15, 16] specified by the values of a binary variable $\xi = \pm 1$. More precisely, let us introduce the following sets

$$\mathcal{I} := \left\{ \mathbf{x}^\mu \in \{-1, 1\}^N \mid 1 \leq \mu \leq p \right\}, \quad \mathcal{X} := \left\{ \boldsymbol{\xi} = \{\xi^\mu\}_{\mu=1}^p \mid \xi^\mu = \pm 1 \right\}, \quad (2)$$

which represent the input set—here we are considering input vectors with binary entries—and the label set, respectively. Depending on the weight vectors $\mathbf{w}$ in $\mathbb{R}^N$, a classical perceptron provides a classification of p input patterns $\{\mathbf{x}^\mu\}_{\mu=1}^p \subset \mathcal{I}$. Given an assigned classification vector $\boldsymbol{\xi} \in \mathcal{X}$ to be implemented, with the chosen activation function the perceptron outputs $\text{sgn}(\mathbf{w} \cdot \mathbf{x}^\mu)$, so that the classical perceptron correctly classifies the input patterns iff

$$\text{sgn}(\mathbf{w} \cdot \mathbf{x}^\mu) = \xi^\mu, \quad \forall \mu = 1, \dots, p. \quad (3)$$

Condition (3) is sometimes rephrased in terms of the so-called *pattern stabilities*, defined as

$$\Delta^\mu := \xi^\mu \mathbf{w} \cdot \mathbf{x}^\mu. \quad (4)$$

Accordingly, the classification of the input pattern is correct—w.r.t. a given chosen target classification $\boldsymbol{\xi}$ —iff the condition

$$\Delta^\mu \geq 0 \quad (5)$$

holds for any $\mu = 1, \dots, p$. In many practical situations, the above condition (5) is strengthened by requiring $\Delta^\mu > \kappa > 0$ for all $\mu = 1, \dots, p$. Indeed, this condition provides a higher stability of the perceptron, as it prevents incorrect classification due to noise in the input pattern components.

2.2. The storage capacity problem

A paramount feature of an artificial neuron is its ability to store and classify input patterns. This performance is typically assessed by the so-called *storage capacity*; namely, by a threshold parameter associated with the volume of input patterns a classical perceptron can correctly classify when their dimension N increases [2, 3].

A convenient statistical way to define and compute the storage capacity of a classical perceptron is by means of Gardner's approach [4–6]; here, one starts by introducing the normalized volume—i.e. the Gardner volume—defined as

$$V_N := \frac{1}{Z_N} \int d\mu(\mathbf{w}) \prod_{\mu=1}^p \Theta(\Delta^\mu - \kappa), \quad (6)$$

where Θ is the Heaviside function, $d\mu(\mathbf{w})$ is the normalized uniform measure on the N -dimensional sphere of radius $\sqrt{N}$:

$$d\mu(\mathbf{w}) = \frac{1}{C_N} \int d\mathbf{w} \delta(\|\mathbf{w}\|^2 - N), \quad C_N = \int d\mathbf{w} \delta(\|\mathbf{w}\|^2 - N) = \frac{2\pi^{N/2} N^{(N-1)/2}}{\Gamma(N/2)}, \quad (7)$$

where $d\mathbf{w} = \prod_{i=1}^N dw_i$ ⁹, and $\Gamma(z)$ is the Euler Gamma function. The Gardner volume V_N quantifies the fraction of weight vectors $\mathbf{w} \in \mathbb{R}^N$ constrained to the sphere of radius $\sqrt{N}$ that correctly classify a set of p input patterns. In the thermodynamic limit $N \rightarrow \infty$, one is interested in the regime where a macroscopic number of patterns can be stored, i.e. $p = \alpha N$ with fixed load $\alpha > 0$. As N increases, the $p = \alpha N$ classification constraints in (6) typically reduce V_N at least exponentially in N . The problem becomes unfeasible when this reduction is super-exponential, in which case V_N vanishes too rapidly to allow storage. The storage capacity is thus defined as the critical value α_c of the load parameter $\alpha = p/N$ that separates the regime where V_N remains exponentially small in N from the regime where it decays super-exponentially [2, 3].

To compute the storage capacity, one can exploit the formal analogy between expression (6) and a statistical mechanics partition function. Then, we will consider the patterns $\mathbf{x}^\mu$ and the labels ξ^μ as independent random variables with independent and identically distributed entries, according to

$$P(x_j^\mu = 1) = P(x_j^\mu = -1) = \frac{1}{2}, \quad P(\xi^\mu = 1) = P(\xi^\mu = -1) = \frac{1}{2}. \quad (8)$$

As a consequence, the volume (6) becomes a random variable whose typical value $V_N \sim e^{N\mathcal{F}}$ is characterized by the free energy [17, 18]

$$\mathcal{F}(\alpha) = \lim_{\substack{p, N \rightarrow \infty \\ p/N = \alpha}} \frac{\langle \ln V_N \rangle_{\mathcal{J}, \mathcal{X}}}{N}. \quad (9)$$

In order to compute the so called quenched average $\langle \ln V_N \rangle_{\mathcal{J}, \mathcal{X}}$, a notoriously difficult task, a most convenient approach is the so-called *replica trick* [17, 19]

$$\langle \ln V_N \rangle_{\mathcal{J}, \mathcal{X}} = \lim_{n \rightarrow 0} \frac{\langle V_N^n \rangle_{\mathcal{J}, \mathcal{X}} - 1}{n} = \lim_{n \rightarrow 0} \frac{\ln \langle V_N^n \rangle_{\mathcal{J}, \mathcal{X}}}{n}, \quad (10)$$

but evaluating $\langle V_N^n \rangle_{\mathcal{J}, \mathcal{X}}$ for n integer.

The replica trick (10) allows us to express the expectation value $\langle \ln V_N \rangle_{\mathcal{J}, \mathcal{X}}$ in terms of the average volume over n replicas of the single perceptron random setting

$$\langle V_N^n \rangle_{\mathcal{J}, \mathcal{X}} = \int \prod_{\gamma=1}^n d\mu(\mathbf{w}_\gamma) \left\langle \prod_{\mu=1}^p \prod_{\gamma=1}^n \Theta(\xi^\mu \mathbf{w}_\gamma \cdot \mathbf{x}^\mu - \kappa) \right\rangle_{\mathcal{J}, \mathcal{X}}, \quad (11)$$

which can be computed in the $N \rightarrow \infty$ limit through a saddle-point approximation. In the replica symmetric scenario, the computation is performed with the introduction of the order parameter q , whose

⁹ Note that, the measure defined as $\lambda(E) := \int_E d\mathbf{w} \delta(\|\mathbf{w}\|^2 - N) \prod_{\mu=1}^p \Theta(\Delta^\mu - \kappa)$, for every Borel subset E of $\mathbb{R}^N$, is a well defined Radon measure on $\mathbb{R}^N$.

value at the saddle point characterizes the typical overlap between two different replicas $\mathbf{w}_1, \mathbf{w}_2$ extracted from the uniform measure within the solution space [18, 20]:

$$q = \left\langle \int d\mu(\mathbf{w}_1) d\mu(\mathbf{w}_2) \frac{\mathbf{w}_1 \cdot \mathbf{w}_2}{N} \prod_{\mu=1}^p \Theta(\xi^\mu \mathbf{w}_1 \cdot \mathbf{x}^\mu - \kappa) \Theta(\xi^\mu \mathbf{w}_2 \cdot \mathbf{x}^\mu - \kappa) \right\rangle. \quad (12)$$

The value of q at the saddle point depends on the value of the load parameter α in (9), and in particular one finds that it monotonically increases in α . Then, the critical value α_c can be characterized as the value of α such that $q \rightarrow 1$, signaling the fact that the typical volume of the solution space shrinks to zero. In the classical perceptron, this procedure yields the final result for the storage capacity in the form [4, 5, 16]

$$\alpha_c(\kappa) = \left(\int_{-\kappa}^{+\infty} \frac{dy}{\sqrt{2\pi}} e^{-y^2/2} (\kappa + y)^2 \right)^{-1}. \quad (13)$$

Interestingly, when $\kappa \rightarrow 0$, one finds $\alpha_c(0) = 2$, which is the value of the storage capacity obtained through a completely different approach, based on a geometric argument due to Cover [21].

3. Quantum storage capacity

Quantum neural networks generalize at the quantum level the notion of a feed-forward neural network. Like their classical counterparts, the fundamental computational unit of a quantum neural network is represented by the so-called quantum perceptron. In recent years, several proposals for implementing a quantum perceptron have been considered [7, 8, 11, 22–35]. Regardless of the specific model, each approach typically involves three essential steps: first, constructing an encoding circuit to map classical input data into a quantum state; second, defining a set of trainable quantum gates—controlled by tunable weights w_{ij} —to realize a quantum counterpart to the non-linear output of the classical perceptron; and finally, setting a readout operation to retrieve a classical output from the quantum system. In this respect, notice that the quantum measurement process itself, in its selective version, amounts to a non-linear operation.

3.1. A discrete model of quantum perceptron

In this work, we focus on a model first proposed in [11]. A perceptron is here implemented by a qubit subjected to an external coupling that operates a unitary transformation parametrized by a classical activation function. Specifically, the k -th qubit in a multi-layered perceptron is acted upon by a unitary transformation $U_k(\mathbf{w}_k, f)$ that depends on a non-linear activation function f and on tunable weights w_{kj} , $j = 1, 2, \dots, k-1$, as follows:

$$U_k(\mathbf{w}_k, f) := \exp \left(-i \arcsin \sqrt{f \left(\sum_{j < k} w_{kj} \sigma_z^{(j)} - b_k \right)} \otimes \sigma_y^{(k)} \right), \quad (14)$$

where $\sigma_\alpha^{(i)}$, $\alpha \in \{x, y, z\}$, denotes the α -Pauli operator acting on the i -th qubit.

Remark 1. The above model of discrete quantum perceptron has been proved to contain the classical perceptron as a limit and thus to provide a universal approximator of continuous functions [11].

In the simplest case, the quantum neural network consists of a single input layer with N qubits, and an output layer with a single qubit. Let $\mathcal{H}_N = \mathbb{C}^{2^N}$ and $|\Psi_N\rangle\langle\Psi_N|$ denote the Hilbert space, and a pure state of the input layer, respectively. Similarly, let $\mathcal{H}_{\text{out}} = \mathbb{C}^2$ and $|\phi\rangle\langle\phi|$ be the Hilbert space and an initial pure state of the output quantum neuron. Without loss of generality, we can choose the initial state of the network in a factorized form, i.e.

$$\rho_{\text{in}} := |\Psi_N\rangle\langle\Psi_N| \otimes |\phi\rangle\langle\phi|. \quad (15)$$

Then, setting the bias $b = 0$ and letting f be the Heaviside Θ function, the unitary action (14) implementing the quantum perceptron reduces to

$$U(\mathbf{w}, \Theta) = \exp \left(-i \arcsin \sqrt{\Theta(\mathbf{w} \cdot \boldsymbol{\sigma}_z)} \otimes \sigma_y^{(\text{out})} \right), \quad (16)$$

where $\sigma_z = (\sigma_z^{(1)}, \sigma_z^{(2)}, \dots, \sigma_z^{(N)})$, $\mathbf{w} = (w_1, \dots, w_N) \in \mathbb{R}^N$, and $\sigma_y^{(\text{out})}$ is the y -Pauli operation performed by the output perceptron. An input vector of binary classical data $\mathbf{x} = (x_1, x_2, \dots, x_N) \in \{\pm 1\}^N$ is naturally encoded as the tensor product of eigen-states of σ_z ; namely,

$$\{\pm 1\}^N \ni \mathbf{x} \mapsto |\mathbf{x}\rangle = |x_1\rangle \otimes |x_2\rangle \otimes \dots \otimes |x_N\rangle, \quad \sigma_z^{(j)} |x_j\rangle = x_j |x_j\rangle. \quad (17)$$

With this notation, the unitary in (16) reads

$$\begin{aligned} U(\mathbf{w}, \Theta) &= \sum_{\mathbf{y} \in \{\pm 1\}^N} |\mathbf{y}\rangle \langle \mathbf{y}| \otimes \exp \left(-i \arcsin \left(\sqrt{\Theta(\mathbf{w} \cdot \mathbf{y})} \right) \sigma_y^{(\text{out})} \right) \\ &= \sum_{\mathbf{y} \in \{\pm 1\}^N} |\mathbf{y}\rangle \langle \mathbf{y}| \otimes \left(\sqrt{1 - \Theta(\mathbf{w} \cdot \mathbf{y})} \mathbb{I}^{(\text{out})} - i \sqrt{\Theta(\mathbf{w} \cdot \mathbf{y})} \sigma_y^{(\text{out})} \right). \end{aligned} \quad (18)$$

By choosing the initial state in (15) as the projector onto $|\mathbf{x}\rangle \otimes |-1\rangle$, where $\sigma_z^{\text{out}} |-1\rangle = -|-1\rangle$, the state of the output qubit becomes

$$\rho_{\mathbf{w}, \mathbf{x}}^{(\text{out})} = \text{tr}_{\mathcal{H}_N} \left(U(\mathbf{w}, \Theta) (|\mathbf{x}\rangle \langle \mathbf{x}| \otimes |-1\rangle \langle -1|) U^\dagger(\mathbf{w}, \Theta) \right) \quad (19)$$

$$= (1 - \Theta(\mathbf{w} \cdot \mathbf{x})) |-1\rangle \langle -1| + \Theta(\mathbf{w} \cdot \mathbf{x}) |1\rangle \langle 1|. \quad (20)$$

The last readout step of the perceptron architecture evaluates the expectation value of σ_z w.r.t. the output state $\rho_{\mathbf{w}, \mathbf{x}}^{(\text{out})}$, yielding

$$\langle \sigma_z \rangle_{\mathbf{w}, \mathbf{x}} = 2\Theta(\mathbf{w} \cdot \mathbf{x}) - 1. \quad (21)$$

Remark 2. The choice of the operator to measure in the readout step is completely arbitrary and should be chosen out of convenience. A measurement in a different basis can always be thought of as an additional fixed unitary step (independent on the data to be classified or the trainable parameter $\mathbf{w}$) just before the measurement is performed.

Remark 3. The discrete quantum perceptron model outlined above is closely related to another proposal put forward in [35] whereby a unitary gate

$$U(z) = \exp \left(i \beta^{(j)}(z) \sigma_y \right), \quad \beta^{(j)}(z) = 2 \arctan \left(\tan^{2^j}(z) \right), \quad (22)$$

is implemented by a so-called ‘repeat until success’ strategy. Namely, $U(z)$ is a rotation around y by an angle $\beta^{(j)}(z)$ that, in the limit for $j \rightarrow \infty$, converges to a step-wise function in the interval $z \in \left[-\frac{\pi}{4}, \frac{\pi}{4} \right]$.

3.2. Quantum storage capacity

In order to compute the storage capacity of the quantum perceptron presented in section 3.1, we follow Gardner’s approach as discussed in section 2.2. We focus upon the volume (6) of those (normalized) weight vectors $\mathbf{w} \in \mathbb{R}^N$ which correctly classify p input patterns $\mathbf{x}^\mu \in \{\pm 1\}^N$, i.e.

$$\begin{aligned} V(\{\mathbf{x}^\mu, \xi^\mu\}_{\mu=1}^p) &= \int d\mu(\mathbf{w}) \prod_{\mu=1}^p \Theta(\xi^\mu \langle \sigma_z \rangle_{\mathbf{w}, \mathbf{x}^\mu}) \\ &= \int d\mu(\mathbf{w}) \prod_{\mu=1}^p \Theta(\xi^\mu (2\Theta(\mathbf{w} \cdot \mathbf{x}^\mu) - 1)), \end{aligned} \quad (23)$$

where, for simplicity, we set the stabilizing threshold $\kappa = 0$. By observing that

$$\Theta(\xi^\mu (2\Theta(\mathbf{w} \cdot \mathbf{x}^\mu) - 1)) = \Theta(\xi^\mu \mathbf{w} \cdot \mathbf{x}^\mu), \quad (24)$$

i.e. the volume (23) is formally equivalent to the Gardner volume (6) of the classical perceptron, we get to the conclusion that it will yield the same limiting value $\alpha_c(0) = 2$.

The above result suggests that, in the quantum perceptron model described in section 3.1, no quantum advantage can be observed at the level of the storage capacity. Notice that the unitary operation in (16) contains the classical non-linear activation function $\Theta(\mathbf{w} \cdot \mathbf{x})$ depending on the rotation angle. However, a non-linear quantum operation is built in the quantum circuit through the selective measurement processes on the output σ_z that extract the eigenvalues ± 1 with which the mean-value $\langle \sigma_z \rangle_{\mathbf{w}, \mathbf{x}}$ is constructed. We will base our subsequent considerations on the quantum non-linear effects

inscribed in the above model (due to the quantum measurement), while eliminating the classical non-linearity (due to the arcsin function). We thus modify the unitary gate of the model in [11] into

$$U(\mathbf{w}, \lambda) := \exp \left(-\frac{i}{2} \frac{\lambda}{\|\mathbf{w}\|} \mathbf{w} \cdot \boldsymbol{\sigma}_z \otimes \sigma_y^{(\text{out})} \right). \quad (25)$$

In the above unitary operator we have also included a *modulation parameter* $\lambda \in [0, +\infty)$ which will allow us to vary the frequency of the oscillations. The removal of non-linearity in the controlled phase of the unitary (25) has the advantage of simplifying the practical implementation with respect to (16). As a matter of fact, the unitary (25) can be even easily implemented on any NISQ device, requiring only two qubits.

When acting on the multiple eigenstates $|\mathbf{x}\rangle$ in (17), the unitary operator $U(\mathbf{w}, \lambda)$ provides a rotation around y by an angle $\mathbf{w} \cdot \mathbf{x}$ yielding the output readout state (see (19))

$$\rho_{\mathbf{w}, \mathbf{x}}^{(\text{out})} = \cos^2 \left(\lambda \frac{\mathbf{w} \cdot \mathbf{x}}{2\|\mathbf{w}\|} \right) |-1\rangle\langle -1| + \sin^2 \left(\lambda \frac{\mathbf{w} \cdot \mathbf{x}}{2\|\mathbf{w}\|} \right) |1\rangle\langle 1| + \frac{1}{2} \sin \left(\frac{\lambda \mathbf{w} \cdot \mathbf{x}}{\|\mathbf{w}\|} \right) (|-1\rangle\langle 1| + |1\rangle\langle -1|). \quad (26)$$

We will consider the measurement of σ_x on the output, whose expectation value on (26) reads

$$\langle \sigma_x \rangle_{\mathbf{w}, \mathbf{x}} = \sin \left(\frac{\lambda \mathbf{w} \cdot \mathbf{x}}{\|\mathbf{w}\|} \right). \quad (27)$$

Compared with (21) we can realize that this outcome corresponds to a (sinusoidally) oscillating activation function with frequency controlled by λ . The choice to measure σ_x rather than σ_z is just a matter of convenience (see remark 2): For input data drawn from the distribution (8), and spherical weights $\mathbf{w}$, the argument of the sin in (27) is centered around zero, which allows one to easily match the desired distribution of the target classifications in (8). Moreover, with this choice the classical limit just corresponds to a linearization of the sin, obtained here through $\lambda \rightarrow 0$. Exchanging σ_x with σ_z in this model would just correspond to the introduction of a fixed bias in the model, with unnecessary complications.

Inserting (27) into (6), provides the following Gardner volume

$$V_N^\lambda \left(\{\mathbf{x}^\mu, \xi^\mu\}_{\mu=1}^p \right) = \int d\mu(\mathbf{w}) \prod_{\mu=1}^p \Theta \left(\xi^\mu \sin \left(\frac{\lambda \mathbf{w} \cdot \mathbf{x}^\mu}{\|\mathbf{w}\|} \right) \right). \quad (28)$$

Following the standard procedure outlined in section 2.2 we find the free energy in the replica symmetric ansatz (see appendix A for the details)

$$\mathcal{F}(\lambda, \alpha) = \lim_{\substack{p, N \rightarrow \infty \\ p/N = \alpha}} \frac{\langle \ln V_N^\lambda \rangle_{\mathcal{J}, \mathcal{R}}}{N} = \lim_{p, N \rightarrow \infty} \lim_{n \rightarrow 0} \frac{\ln \langle (V_N^\lambda)^n \rangle_{\mathcal{J}, \mathcal{R}}}{nN} = \text{extr}_{0 \leq q < 1} \mathcal{G}(\lambda, \alpha, q), \quad (29)$$

where

$$\mathcal{G}(\lambda, \alpha, q) := \alpha \int_{\mathbb{R}} D\omega \ln \Psi(\lambda, q, \omega) + \frac{1}{2} \left[\frac{q}{1-q} + \ln(1-q) \right]. \quad (30)$$

The saddle point equation $\partial_q \mathcal{G}(\lambda, \alpha, q) = 0$ then yields a relation between the load parameter α and the typical overlap q :

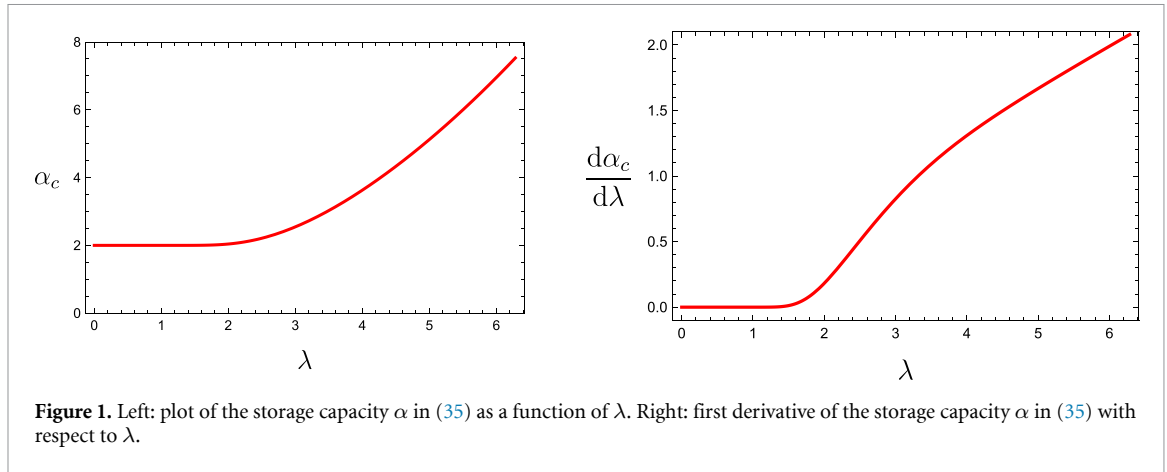
$$\begin{aligned} \alpha(\lambda, q) &= -\frac{q}{2(1-q)^2} \left(\int \frac{d\omega}{\sqrt{2\pi}} e^{-\frac{\omega^2}{2}} \partial_q \ln \Psi(\lambda, q, \omega) \right)^{-1} \\ &= -\frac{q}{2(1-q)^2} \left(\int \frac{d\omega}{\sqrt{2\pi}} e^{-\frac{\omega^2}{2}} \frac{\Phi(\lambda, q, \omega)}{\Psi(\lambda, q, \omega)} \right)^{-1}, \end{aligned} \quad (31)$$

where we set

$$\Psi(\lambda, q, \omega) := \sum_{k \in \mathbb{Z}} \int_{\epsilon_1(\lambda, q, \omega, k)}^{\epsilon_2(\lambda, q, \omega, k)} \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad \Phi(\lambda, q, \omega) := \partial_q \Psi(\lambda, q, \omega), \quad (32)$$

with $\epsilon_1(\lambda, q, \omega, k)$ and $\epsilon_2(\lambda, q, \omega, k)$ given by

$$\epsilon_1(\lambda, q, \omega, k) := \frac{-\lambda\sqrt{q}\omega + 2k\pi}{\lambda\sqrt{1-q}}, \quad \epsilon_2(\lambda, q, \omega, k) := \frac{-\lambda\sqrt{q}\omega + (2k+1)\pi}{\lambda\sqrt{1-q}}. \quad (33)$$



The critical storage capacity $\alpha_c(\lambda)$ is then obtained via the limit

$$\alpha_c(\lambda) = \lim_{q \rightarrow 1^-} \alpha(\lambda, q). \quad (34)$$

In appendix B it is shown that the limit in (34) can be analytically computed, resulting in the λ -dependent storage capacity

$$\alpha_c(\lambda) = \left(\sum_{k=0}^{+\infty} \int_0^{2\pi/\lambda} \frac{d\omega}{\sqrt{2\pi}} \omega^2 \exp \left(-\frac{1}{2} \left(\omega + \frac{2\pi}{\lambda} k \right)^2 \right) \right)^{-1}. \quad (35)$$

Notice that the series is uniformly convergent in $\lambda > 0$ (see appendix C) then, when $\lambda \rightarrow 0^+$, only the contribution $k=0$ survives, so that

$$\alpha_c(0) = \lim_{\lambda \rightarrow 0^+} \alpha_c(\lambda) = \left(\int_0^{+\infty} \frac{d\omega}{\sqrt{2\pi}} \omega^2 e^{-\frac{\omega^2}{2}} \right)^{-1} = 2, \quad (36)$$

and one retrieves the standard storage capacity. Furthermore, the function $\alpha_c(\lambda)$ is finite and infinitely differentiable at $\lambda=0$; however, all derivatives vanish at that point so that the storage capacity is not analytic there. On the other hand, $\lim_{\lambda \rightarrow +\infty} \alpha_c(\lambda) = +\infty$ (see also appendix D).

In all cases, $\alpha_c(\lambda) \geq 2$. Indeed, changing the integration variable to $t = \frac{\omega}{\sqrt{2}} + \frac{\sqrt{2}\pi}{\lambda} k$, one obtains

$$\begin{aligned} \sum_{k=0}^{+\infty} \int_0^{2\pi/\lambda} \frac{d\omega}{\sqrt{2\pi}} \omega^2 \exp \left(-\frac{1}{2} \left(\omega + \frac{2\pi}{\lambda} k \right)^2 \right) &= 2 \sum_{k=0}^{+\infty} \int_{\sqrt{2}\pi k/\lambda}^{\sqrt{2}\pi(k+1)/\lambda} \frac{dt}{\sqrt{\pi}} \left(t - \frac{\sqrt{2}\pi k}{\lambda} \right)^2 e^{-t^2} \\ &= 2 \int_0^{+\infty} \frac{dt}{\sqrt{\pi}} t^2 e^{-t^2} + 2 \sum_{k=0}^{+\infty} \int_{\sqrt{2}\pi k/\lambda}^{\sqrt{2}\pi(k+1)/\lambda} \frac{dt}{\sqrt{\pi}} \left(\frac{2\pi^2 k^2}{\lambda^2} - \frac{2\sqrt{2}\pi t}{\lambda} \right) e^{-t^2} \\ &= \frac{1}{2} + 2 \sum_{k=0}^{+\infty} \int_{\sqrt{2}\pi k/\lambda}^{\sqrt{2}\pi(k+1)/\lambda} \frac{dt}{\sqrt{\pi}} \frac{2\pi k}{\lambda} \left(\frac{\pi k}{\lambda} - \sqrt{2} t \right) e^{-t^2} \leq \frac{1}{2}, \end{aligned} \quad (37)$$

where in the last line we used the fact that for $t \geq \sqrt{2}\pi k/\lambda$ one has

$$\frac{\pi k}{\lambda} - \sqrt{2} t \leq -\frac{\pi k}{\lambda} \leq 0. \quad (38)$$

Inserting the bound (37) into (35) one concludes that

$$\alpha_c(\lambda) \geq 2 \quad \text{for all } \lambda > 0. \quad (39)$$

The left panel of figure 1 illustrates the behavior of α in (35) as a function of λ . The value of α remains nearly constant at value 2 until λ approaches 2, beyond which it increases monotonically and without bound. The right panel, instead, presents the derivative of α with respect to λ , plotted against λ . The derivative remains close to zero up to $\lambda \approx 1.6$, after which it begins to increase gradually and monotonically. An evidence of the fact that α increases nonlinearly with λ .

4. Conclusions and outlook

Summarizing, building on the discrete model presented in [11] for implementing a quantum perceptron, we modified the unitary gate (25) to explore how variations in the oscillation period affect the system's behavior. Then, using the replica method, we analytically computed the storage capacity as a function of the oscillation frequency λ . In the limit of vanishing frequency, the classical value of 2 is recovered. However, as the frequency increases, we observe an indefinite enhancement in the quantum storage capacity beyond the classical threshold. This behavior is as expected taking into account that the periodicity of the activation function provides a 'finer partitioning' of the input-state space, thereby leading to an increased perceptron storage capacity. However, as the frequency rises, this advantage may turn into a liability, since overfitting tends to take over. This kind of drawback can be ascertained via a thorough study of the generalization error in a teacher-student setting, where explicit overfitting effects are expected to emerge and to be accessible by means of Gardner's approach. We plan to address this intriguing problem in a future work together with the robustness of our results in the presence of noise in the unitary operations. This kind of noise can be mimicked by means of either normal distributed random weights or inputs in a statistical learning scenario [2]. Furthermore, it is also important to explore how these results are affected by a replica symmetry breaking, for example within the one-step replica symmetric ansatz.

As a general remark, the observed improvement of the storage capacity beyond the classical threshold $\alpha = 2$ arises solely from the specific form of the activation function. This indicates that, in principle, comparable enhancements could be achieved within a fully classical framework. Supporting evidence appears in [36], where the authors argue that non-monotonic activation functions can substantially improve neural-network performance. In particular, they show that employing a sine activation function in a feedforward architecture enhances both generalization and learning speed relative to sigmoid activations, thereby providing a classical benchmark for our findings. Actually, in the quantum perceptron model considered here, the nonlinearity of the activation function does not stem from an intrinsic quantum mechanism. Instead, it emerges from taking expectation values of a measured observable—a procedure that is effectively classical. In this sense, the nonlinearity is 'measurement-induced' through averaging and could, in principle, be reproduced within a classical setting. By contrast, a genuinely quantum source of nonlinearity could arise, for example, from a selective measurement process in which the perceptron's output is determined directly by a single measurement outcome rather than by an averaged expectation value. Such a mechanism would introduce a form of nonlinearity that cannot be reproduced classically. Another potential route to quantum advantage may emerge at the level of a full neural network, where interference between different perceptrons becomes relevant and the resulting performance gains stem from intrinsically quantum interference effects.

Acknowledgments

The authors acknowledge financial support from the PNRR MUR Project PE0000023-NQSTI and from the PRIN MUR Project 2022SW3RPY. GG is partially supported by Istituto Nazionale di Fisica Nucleare (INFN) through the Project 'QUANTUM' and by the Italian National Group of Mathematical Physics (GNFM-INDAM). GG acknowledges financial support from the University of Bari through the 2023-UNBACLE-0245516 Grant.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Author contributions

Fabio Benatti  0000-0002-0712-2057

Conceptualization (equal), Data curation (equal), Formal analysis (equal), Funding acquisition (equal), Investigation (equal), Methodology (equal), Project administration (equal), Resources (equal), Software (equal), Supervision (equal), Validation (equal), Visualization (equal), Writing – original draft (equal), Writing – review & editing (equal)

Masoud Gharahi  0000-0002-7515-6179

Conceptualization (equal), Data curation (equal), Formal analysis (equal), Funding acquisition (equal), Investigation (equal), Methodology (equal), Project administration (equal), Resources (equal), Software (equal), Supervision (equal), Validation (equal), Visualization (equal), Writing – original draft (equal), Writing – review & editing (equal)

Giovanni Gramegna  0000-0001-7532-1704

Conceptualization (equal), Data curation (equal), Formal analysis (equal), Funding acquisition (equal), Investigation (equal), Methodology (equal), Project administration (equal), Resources (equal), Software (equal), Supervision (equal), Validation (equal), Visualization (equal), Writing – original draft (equal), Writing – review & editing (equal)

Stefano Mancini  0000-0002-3797-3987

Conceptualization (equal), Data curation (equal), Formal analysis (equal), Funding acquisition (equal), Investigation (equal), Methodology (equal), Project administration (equal), Resources (equal), Software (equal), Supervision (equal), Validation (equal), Visualization (equal), Writing – original draft (equal), Writing – review & editing (equal)

Vincenzo Parisi  0000-0001-9563-6471

Conceptualization (equal), Data curation (equal), Formal analysis (equal), Funding acquisition (equal), Investigation (equal), Methodology (equal), Project administration (equal), Resources (equal), Software (equal), Supervision (equal), Validation (equal), Visualization (equal), Writing – original draft (equal), Writing – review & editing (equal)

Appendix A. Computation of the quantum storage capacity (34)

We now compute the expression of the storage capacity $\alpha_c(\lambda)$ of Subsection 3.2. We consider a data-set of the form $\{\mathbf{x}^\mu, \xi^\mu\}_{\mu=1}^p$, where $\mathbf{x}^\mu \in \{-1, 1\}^N$, while $\xi^\mu = \pm 1$ is a binary label. Using the parity of $\sin(z)$, the Gardner volume (28) can be rewritten as

$$V_N^\lambda(\{\mathbf{x}^\mu, \xi^\mu\}_{\mu=1}^p) = \int d\mu(\mathbf{w}) \prod_{\mu=1}^p \Theta \left(\sin \left(\lambda \frac{\mathbf{w} \cdot \mathbf{r}^\mu}{\sqrt{N}} \right) \right), \quad \mathbf{r}^\mu = \xi^\mu \mathbf{x}^\mu. \quad (40)$$

The quantity of interest is the expectation value of $\ln V$ w.r.t. the distribution of patterns and the labels. Note that the distribution (8) on $\{\mathbf{x}^\mu, \xi^\mu\}_{\mu=1}^p$ induces the distribution

$$P(r_j^\mu = 1) = P(r_j^\mu = -1) = \frac{1}{2}. \quad (41)$$

In the following we will denote with $\langle \cdot \rangle$ expectations with respect to the distribution (41). The computation of $\langle \ln V_N^\lambda \rangle$ is performed by the replica trick (10), which leads us to consider the following expectation value

$$\langle (V_N^\lambda)^n \rangle = \int \prod_{\gamma=1}^n d\mu(\mathbf{w}_\gamma) \prod_{\mu=1}^p \left\langle \prod_{\gamma=1}^n \Theta \left(\sin \left(\lambda \frac{\mathbf{w}_\gamma \cdot \mathbf{r}^\mu}{\sqrt{N}} \right) \right) \right\rangle. \quad (42)$$

Using the spin glass order parameters

$$q_{\gamma\delta} := \frac{1}{N} \mathbf{w}_\gamma \cdot \mathbf{w}_\delta, \quad (43)$$

the integral (42) can be rewritten as

$$\langle (V_N^\lambda)^n \rangle = \frac{1}{C_N^n} \int dQ \int d\mathbf{W} \delta(\mathbf{W}^T \mathbf{W} - NQ) \prod_{\mu=1}^p \left\langle \prod_{\gamma=1}^n \Theta \left(\sin \left(\lambda \frac{\mathbf{w}_\gamma \cdot \mathbf{r}^\mu}{\sqrt{N}} \right) \right) \right\rangle, \quad (44)$$

where, we introduced the $N \times n$ matrix $\mathbf{W} = (\mathbf{w}_1, \dots, \mathbf{w}_n)$ whose columns are $\mathbf{w}_\gamma$, the $n \times n$ matrix Q whose elements are $q_{\gamma\delta}$ and we introduced the notations

$$d\mathbf{W} := \prod_{\gamma=1}^n d\mathbf{w}_\gamma, \quad dQ := \prod_{\gamma < \delta} dq_{\gamma\delta}, \quad (45)$$

$$\delta(\mathbf{W}^T \mathbf{W} - NQ) := \prod_{\gamma \leq \delta} \delta(\mathbf{w}_\gamma \cdot \mathbf{w}_\delta - NQ). \quad (46)$$

Note that the integration in Q is done only over the upper off-diagonal terms, since the matrix is symmetric by definition and the diagonal terms are fixed to $q_{\gamma\gamma} = 1$ by the normalization condition $\mathbf{w}_\gamma \cdot \mathbf{w}_\gamma = N$. A standard computation (see appendix A.2 for the details) shows that to the leading order in the thermodynamic limit $N \rightarrow \infty$:

$$\prod_{\mu=1}^p \left\langle \prod_{\gamma=1}^n \Theta \left(\sin \left(\lambda \frac{\mathbf{w}_\gamma \cdot \mathbf{r}^\mu}{\sqrt{N}} \right) \right) \right\rangle \delta(\mathbf{W}^T \mathbf{W} - NQ) \simeq e^{NE_\lambda(Q)} \delta(\mathbf{W}^T \mathbf{W} - NQ), \quad (47)$$

with

$$E_\lambda(Q) = \alpha \ln \left(\frac{1}{(2\pi)^n} \int_{\Sigma^n} d\mathbf{z} \int_{\mathbb{R}^n} d\mathbf{y} e^{i\mathbf{z} \cdot \mathbf{y} - \frac{\lambda^2}{2} \mathbf{y} \cdot Q \mathbf{y}} \right), \quad (48)$$

where $\mathbf{z} := (z_1, \dots, z_n)$, $\mathbf{y} := (y_1, \dots, y_n)$ and

$$\Sigma := \{z \in \mathbb{R} \mid 2k\pi \leq z \leq (2k+1)\pi, k \in \mathbb{Z}\}. \quad (49)$$

In other words, in the thermodynamic limit only the overlap matrix Q is relevant, while the remaining degrees of freedom in $\mathbf{W}$ are redundant. This redundancy can be integrated out, considering

$$e^{NS(Q)} = \int d\mathbf{W} \delta(\mathbf{W}^T \mathbf{W} - QN), \quad (50)$$

which represents the volume in the $\mathbf{W}$ space consistent with the constraints (43). The details of this change of variables, including the evaluation of the Jacobian determinant involved, are worked out explicitly in [37]. The result, up to irrelevant constants, is given by

$$e^{NS(Q)} \propto \det(Q)^{\frac{N-n-1}{2}}. \quad (51)$$

Therefore, to the leading order in $N \rightarrow \infty$:

$$S(Q) \simeq \frac{1}{2} \ln \det(Q). \quad (52)$$

Using (47) and (50), equation (44) can be recast as

$$\langle (V_N^\lambda)^n \rangle = \int dQ e^{N[E_\lambda(Q) + S(Q)]}. \quad (53)$$

Note that $E_\lambda(Q)$ is the only term containing the information on the specific problem we are considering (through the integration domain Σ), while the term $S(Q)$ is purely geometric in nature. Therefore they are usually called the energetic and entropic contribution, respectively. Expression (53) is well suited for a computation of the replicated volume in the thermodynamic limit $N \rightarrow \infty$ through a saddle-point approximation:

$$\frac{\ln \langle (V_N^\lambda)^n \rangle}{N} = \text{extr}_Q [E_\lambda(Q) + S(Q)]. \quad (54)$$

A.1. Replica symmetric ansatz

In the replica symmetric ansatz, namely $q_{\gamma\delta} = q$ for each $\gamma \neq \delta$, (48) takes the form

$$E_\lambda(Q) = \alpha \ln \left(\frac{1}{(2\pi)^n} \int_{\Sigma^n} d\mathbf{z} \int_{\mathbb{R}^n} d\mathbf{y} e^{i\mathbf{z} \cdot \mathbf{y} - \frac{\lambda^2}{2} (1-q) \|\mathbf{y}\|^2 - \frac{\lambda^2}{2} q (\sum_{\gamma=1}^n y_\gamma)^2} \right). \quad (55)$$

The last exponential can be dealt with the introduction of an additional auxiliary Gaussian variable through the Hubbard-Stratonovich transformation

$$e^{-\frac{q}{2} \xi^2} = \int_{\mathbb{R}} D\omega e^{-i\sqrt{q}\xi\omega}, \quad \text{where} \quad D\omega := d\omega \frac{e^{-\frac{\omega^2}{2}}}{\sqrt{2\pi}}, \quad (56)$$

which yields factorization over the replicas:

$$\begin{aligned} E_\lambda(Q) &= \alpha \ln \left(\int_{\mathbb{R}} D\omega \left(\frac{1}{2\pi} \int_{\Sigma} dz \int_{\mathbb{R}} dy e^{i(z - \sqrt{q}\lambda)y - \frac{\lambda^2}{2}(1-q)y^2} \right)^n \right) \\ &= \alpha \ln \left(\int_{\mathbb{R}} D\omega \left(\frac{1}{[2\pi \lambda^2 (1-q)]^{1/2}} \int_{\Sigma} dz e^{-\frac{(z - \sqrt{q}\lambda)^2}{2\lambda^2(1-q)}} \right)^n \right). \end{aligned} \quad (57)$$

After a change of variables, we get to the expression

$$E_\lambda(Q) = \alpha \ln \left(\int_{\mathbb{R}} D\omega \left(\int_{\Sigma_\lambda^q} Dz \right)^n \right), \quad (58)$$

where

$$\Sigma_\lambda^q := \bigcup_{k \in \mathbb{Z}} \left[\frac{-\lambda\sqrt{q}\omega + 2k\pi}{\lambda\sqrt{1-q}}, \frac{-\lambda\sqrt{q}\omega + (2k+1)\pi}{\lambda\sqrt{1-q}} \right]. \quad (59)$$

The leading order of (58) in the $n \rightarrow 0$ limit is given by

$$E_\lambda(Q) \simeq n\alpha \int_{\mathbb{R}} D\omega \ln \left(\int_{\Sigma_\lambda^q} Dz \right) = n\alpha \int_{\mathbb{R}} D\omega \ln \Psi(\lambda, q, \omega) \quad (60)$$

where we set

$$\Psi(\lambda, q, \omega) := \sum_{k \in \mathbb{Z}} \int_{\epsilon_1(\lambda, q, \omega, k)}^{\epsilon_2(\lambda, q, \omega, k)} \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}, \quad (61)$$

with

$$\epsilon_1(\lambda, q, \omega, k) := \frac{-\lambda\sqrt{q}\omega + 2k\pi}{\lambda\sqrt{1-q}}, \quad \epsilon_2(\lambda, q, \omega, k) := \frac{-\lambda\sqrt{q}\omega + (2k+1)\pi}{\lambda\sqrt{1-q}}. \quad (62)$$

To evaluate the entropic term within the replica symmetric ansatz, it is useful to note that the eigenvalue spectrum of the overlap matrix Q can be obtained explicitly: $n-1$ eigenvalues are equal to $(1-q)$, and a single eigenvalue is $(1+q(n-1))$. The determinant of Q thus reads

$$\det Q = (1+q(n-1))(1-q)^{n-1}. \quad (63)$$

Consequently, to the leading order in $n \rightarrow 0$:

$$S(Q) = \frac{1}{2} \ln \det Q \quad (64)$$

$$= \frac{1}{2} [\ln(1+q(n-1)) + (n-1) \ln(1-q)] \quad (65)$$

$$= \frac{n}{2} \left[\frac{q}{1-q} + \ln(1-q) \right] + O(n^2). \quad (66)$$

Therefore, in the replica symmetric ansatz one has

$$\frac{\ln \langle (V_N^\lambda)^n \rangle}{nN} = \text{extr}_{0 \leq q \leq 1} \mathcal{G}(\lambda, \alpha, q), \quad (67)$$

where

$$\mathcal{G}(\lambda, \alpha, q) := \alpha \int_{\mathbb{R}} D\omega \ln \Psi(\lambda, q, \omega) + \frac{1}{2} \left[\frac{q}{1-q} + \ln(1-q) \right]. \quad (68)$$

The saddle-point equation $\partial_q \mathcal{G}(\lambda, \alpha, q) = 0$ then reads

$$\alpha \int_{\mathbb{R}} D\omega \frac{\partial}{\partial q} \ln \Psi(\lambda, q, \omega) + \frac{q}{2(1-q)^2} = 0, \quad (69)$$

which finally yields

$$\alpha(\lambda, q) = -\frac{q}{2(1-q)^2} \left(\int \frac{d\omega}{\sqrt{2\pi}} e^{-\frac{\omega^2}{2}} \frac{\Phi(\lambda, q, \omega)}{\Psi(\lambda, q, \omega)} \right)^{-1}, \quad (70)$$

where we wrote explicitly the gaussian measure $D\omega$ and we set

$$\Phi(\lambda, q, \omega) := \partial_q \Psi(\lambda, q, \omega). \quad (71)$$

A.2. Derivation of $E_\lambda(Q)$

In this section we present the detailed derivation of equations (47) and (48).

First, note that for $\eta \in \mathbb{R}$

$$\Theta(\sin(\eta)) = \begin{cases} 1 & \text{if } \eta \in \Sigma \\ 0 & \text{otherwise} \end{cases}, \quad \Sigma = \{\eta \in \mathbb{R} \mid 2k\pi \leq \eta \leq (2k+1)\pi, k \in \mathbb{Z}\} \quad (72)$$

admits the integral representation

$$\Theta(\sin(\eta)) = \int_{\Sigma} dz \delta(z - \eta) = \frac{1}{2\pi} \int_{\Sigma} dz \int_{\mathbb{R}} dy e^{iy(z-\eta)}. \quad (73)$$

Therefore:

$$\begin{aligned} & \prod_{\mu=1}^p \left\langle \prod_{\gamma=1}^n \Theta \left(\sin \left(\lambda \frac{\mathbf{w}_\gamma \cdot \mathbf{r}^\mu}{\sqrt{N}} \right) \right) \right\rangle \\ &= \int_{\Sigma^{np}} dZ \int_{\mathbb{R}^{np}} dY \frac{1}{(2\pi)^{np}} e^{iZ \cdot Y} \prod_{\mu=1}^p \left\langle \exp \left(-i \frac{1}{\sqrt{N}} \sum_{\gamma=1}^n y_\gamma^\mu \mathbf{w}_\gamma \cdot \mathbf{r}^\mu \right) \right\rangle, \end{aligned} \quad (74)$$

where we introduced the short-hand notation $Z := (z_z^\mu)_{z,\mu}$, $Y := (y_\gamma^\mu)_{z,\mu}$. Using the fact that the components of $\mathbf{r}^\mu$ are independent and distributed according to (41), we get

$$\begin{aligned} \left\langle \exp \left(-i \frac{1}{\sqrt{N}} \sum_{\gamma=1}^n y_\gamma^\mu \mathbf{w}_\gamma \cdot \mathbf{r}^\mu \right) \right\rangle &= \prod_{j=1}^N \left\langle \exp \left(-i \frac{1}{\sqrt{N}} \sum_{\gamma=1}^n y_\gamma^\mu w_{\gamma j} r_j^\mu \right) \right\rangle \\ &= \prod_{j=1}^N \cos \left(\frac{1}{\sqrt{N}} \sum_{\gamma=1}^n y_\gamma^\mu w_{\gamma j} \right). \end{aligned} \quad (75)$$

Now note that to the leading order in $N \rightarrow \infty$:

$$\begin{aligned} \prod_{j=1}^N \cos \left(\frac{1}{\sqrt{N}} \sum_{\gamma=1}^n y_\gamma^\mu w_{\gamma j} \right) &= \exp \left(\sum_{j=1}^N \ln \left[\cos \left(\frac{1}{\sqrt{N}} \sum_{\gamma=1}^n y_\gamma^\mu w_{\gamma j} \right) \right] \right) \\ &\simeq \exp \left(-\frac{1}{2N} \sum_{j=1}^N \sum_{\gamma, \delta=1}^n y_\gamma^\mu y_\delta^\mu w_{\gamma j} w_{\delta j} \right) \\ &= \exp \left(-\frac{1}{2} \sum_{\gamma, \delta=1}^n y_\gamma^\mu y_\delta^\mu q_{\gamma \delta} \right), \end{aligned} \quad (76)$$

where we used the expansion $\ln(\cos(\varepsilon)) \simeq -\varepsilon^2/2$ as $\varepsilon \rightarrow 0$ and the definition (43). Insertion of (76) into (75) allows to rewrite (74) as

$$\begin{aligned} & \prod_{\mu=1}^p \left\langle \prod_{\gamma=1}^n \Theta \left(\sin \left(\lambda \frac{\mathbf{w}_\gamma \cdot \mathbf{r}^\mu}{\sqrt{N}} \right) \right) \right\rangle \\ &= \int_{\Sigma^{np}} dZ \int_{\mathbb{R}^{np}} dY \frac{1}{(2\pi)^{np}} \prod_{\mu=1}^p \exp \left(i \sum_{\gamma=1}^n z_\gamma^\mu y_\gamma^\mu - \frac{1}{2} \sum_{\gamma, \delta=1}^n y_\gamma^\mu y_\delta^\mu q_{\gamma \delta} \right) \end{aligned}$$

$$\begin{aligned}
&= \left(\frac{1}{(2\pi)^n} \int_{\Sigma^n} dz \int_{\mathbb{R}^n} dy e^{iz \cdot y - \frac{\lambda^2}{2} y \cdot Qy} \right)^p \\
&= e^{NE_\lambda(Q)},
\end{aligned} \tag{77}$$

with $E_\lambda(Q)$ given in (48).

Appendix B. Computation of $\lim_{q \rightarrow 1^-} \alpha(\lambda, q)$

In order to handle the function $\Psi(\lambda, q, \omega)$ in (32), we write

$$e^{-\frac{z^2}{2}} = \int_{-\infty}^{+\infty} dy \frac{e^{iyz}}{\sqrt{2\pi}} e^{-y^2/2}; \tag{78}$$

then, integrating

$$\begin{aligned}
\int_{\epsilon_1(\lambda, q, \omega, k)}^{\epsilon_2(\lambda, q, \omega, k)} \frac{dz}{\sqrt{2\pi}} \frac{e^{iyz}}{\sqrt{2\pi}} &= \frac{1}{2\pi iy} \left(e^{iy \frac{(2k+1)\pi - \lambda\sqrt{q}\omega}{\lambda\sqrt{1-q}}} - e^{iy \frac{2k\pi - \lambda\sqrt{q}\omega}{\lambda\sqrt{1-q}}} \right) \\
&= e^{iy \frac{2k\pi}{\lambda\sqrt{1-q}}} e^{-iy \frac{\sqrt{q}\omega}{\sqrt{1-q}}} \frac{e^{\frac{iy\pi}{\lambda\sqrt{1-q}}} - 1}{2\pi iy}.
\end{aligned} \tag{79}$$

The partial sums over $-n \leq k \leq n$ provide a sequence of *Dirichlet kernels*

$$\sum_{k=-n}^n e^{iy \frac{2k\pi}{\lambda\sqrt{1-q}}} = \frac{\sin\left(\frac{2\pi y}{\lambda\sqrt{1-q}} \left(n + \frac{1}{2}\right)\right)}{\sin\left(\frac{\pi y}{\lambda\sqrt{1-q}}\right)} =: D_n\left(\frac{2\pi y}{\lambda\sqrt{1-q}}\right). \tag{80}$$

Sending $n \rightarrow +\infty$, one gets the so-called *Dirac comb* distribution:

$$\lim_{n \rightarrow +\infty} D_n\left(\frac{2\pi y}{\lambda\sqrt{1-q}}\right) = 2\pi \sum_{k=-\infty}^{+\infty} \delta\left(\frac{2\pi y}{\lambda\sqrt{1-q}} - 2k\pi\right) = \lambda\sqrt{1-q} \sum_{k=-\infty}^{+\infty} \delta\left(y - \lambda k\sqrt{1-q}\right), \tag{81}$$

which, inserted into (32) finally yields

$$\begin{aligned}
\Psi(\lambda, q, \omega) &= \frac{1}{2} + \sum_{k=1}^{+\infty} \frac{e^{-(1-q)\lambda^2 k^2/2}}{\pi k} \left(1 - (-1)^k\right) \sin(k\lambda\sqrt{q}\omega) \\
&= \frac{1}{2} + \frac{2}{\pi} \sum_{m=0}^{+\infty} \frac{e^{-(1-q)\lambda^2 (2m+1)^2/2}}{2m+1} \sin((2m+1)\lambda\sqrt{q}\omega).
\end{aligned} \tag{82}$$

To compute the critical storage capacity as a function of λ , we need to compute the limit

$$\alpha(\lambda) = \lim_{q \rightarrow 1^-} \alpha(\lambda, q) = -\frac{1}{2} \left(\lim_{q \rightarrow 1^-} (1-q)^2 I(\lambda, q) \right)^{-1}, \tag{83}$$

where we set

$$I(\lambda, q) := \int_{-\infty}^{+\infty} \frac{d\omega}{\sqrt{2\pi}} e^{-\frac{\omega^2}{2}} \frac{\partial_q \Psi(\lambda, q, \omega)}{\Psi(\lambda, q, \omega)}. \tag{84}$$

Let us consider

$$J(\lambda, q) := (1-q)^2 I(\lambda, q) = (1-q)^2 \int_{-\infty}^{+\infty} \frac{d\omega}{\sqrt{2\pi}} e^{-\frac{\omega^2}{2}} \frac{\partial_q \Psi(\lambda, q, \omega)}{\Psi(\lambda, q, \omega)}. \tag{85}$$

By differentiating (82) term by term, we get

$$\begin{aligned}
\Phi(\lambda, q, \omega) := \partial_q \Psi(\lambda, q, \omega) &= \frac{1}{\pi} \sum_{m=0}^{+\infty} e^{-(1-q)\lambda^2 (2m+1)^2/2} \left(\lambda^2 (2m+1) \sin((2m+1)\lambda\sqrt{q}\omega) \right. \\
&\quad \left. + \frac{\lambda\omega}{\sqrt{q}} \cos((2m+1)\lambda\sqrt{q}\omega) \right).
\end{aligned} \tag{86}$$

In order to deal with $J(\lambda, q)$ in (85), notice that, for all integers $n \in \mathbb{Z}$, one has

$$\Psi(\lambda, q, \omega \pm nT) = \Psi(\lambda, q, \omega), \quad (87)$$

$$\Phi(\lambda, q, \omega \pm nT) = \Phi(\lambda, q, \omega) \pm n\mathcal{X}(\lambda, q, \omega), \quad (88)$$

where $T := \frac{2\pi}{\lambda\sqrt{q}}$, and

$$\mathcal{X}(\lambda, q, \omega) := \frac{2}{q} \sum_{m=0}^{+\infty} e^{-(1-q)\lambda^2(2m+1)^2/2} \cos((2m+1)\lambda\sqrt{q}\omega), \quad (89)$$

$$\mathcal{X}(\lambda, q, \omega \pm nT) = \mathcal{X}(\lambda, q, \omega). \quad (90)$$

One then rewrites

$$\begin{aligned} J(\lambda, q) &= (1-q)^2 \sum_{n=0}^{+\infty} \left(\int_{nT}^{(n+1)T} + \int_{-(n+1)T}^{-nT} \right) \frac{d\omega}{\sqrt{2\pi}} e^{-\frac{\omega^2}{2}} \frac{\Phi(\lambda, q, \omega)}{\Psi(\lambda, q, \omega)} \\ &= (1-q)^2 \sum_{n=0}^{+\infty} \left(\int_0^T \frac{d\omega}{\sqrt{2\pi}} e^{-(\omega+nT)^2/2} \frac{\Phi(\lambda, q, \omega) + n\mathcal{X}(\lambda, q, \omega)}{\Psi(\lambda, q, \omega)} \right. \\ &\quad \left. + \int_{-T}^0 \frac{d\omega}{\sqrt{2\pi}} e^{-(\omega-nT)^2/2} \frac{\Phi(\lambda, q, \omega) - n\mathcal{X}(\lambda, q, \omega)}{\Psi(\lambda, q, \omega)} \right). \end{aligned} \quad (91)$$

Changing ω into $-\omega$ in the second integral yields

$$\begin{aligned} J(\lambda, q) &= (1-q)^2 \sum_{n=0}^{+\infty} \int_0^T \frac{d\omega}{\sqrt{2\pi}} e^{-(\omega+nT)^2/2} \\ &\quad \times \left[\frac{\Phi(\lambda, q, \omega)}{\Psi(\lambda, q, \omega)} + \frac{\Phi(\lambda, q, -\omega)}{\Psi(\lambda, q, -\omega)} + n\mathcal{X}(\lambda, q, \omega) \left(\frac{1}{\Psi(\lambda, q, \omega)} - \frac{1}{\Psi(\lambda, q, -\omega)} \right) \right]. \end{aligned} \quad (92)$$

Now, by changing the integration variable from ω to $\nu = \omega/\sqrt{1-q}$, we get

$$\begin{aligned} J(\lambda, q) &= (1-q)^{5/2} \sum_{n=0}^{+\infty} \int_0^{2\pi/(\lambda\sqrt{q(1-q)})} \frac{d\nu}{\sqrt{2\pi}} e^{-\frac{1}{2}(\nu\sqrt{1-q}+2\pi n/(\lambda\sqrt{q}))^2} \\ &\quad \times \left[\frac{\Phi(\lambda, q, \nu\sqrt{1-q})}{\Psi(\lambda, q, \nu\sqrt{1-q})} + \frac{\Phi(\lambda, q, -\nu\sqrt{1-q})}{\Psi(\lambda, q, -\nu\sqrt{1-q})} \right. \\ &\quad \left. + n\mathcal{X}(\lambda, q, \nu\sqrt{1-q}) \left(\frac{1}{\Psi(\lambda, q, \nu\sqrt{1-q})} - \frac{1}{\Psi(\lambda, q, -\nu\sqrt{1-q})} \right) \right]. \end{aligned} \quad (93)$$

Setting $x_m := (2m+1)\sqrt{1-q}$, so that $\Delta x_m = x_{m+1} - x_m = 2\sqrt{1-q}$, we can recast

$$\Psi(\lambda, q, \nu\sqrt{1-q}) = \frac{1}{2} + \frac{1}{\pi} \sum_{m=0}^{+\infty} \Delta x_m \frac{e^{-\lambda^2 x_m^2/2}}{x_m} \sin(x_m \lambda \nu \sqrt{q}), \quad (94)$$

$$\begin{aligned} \Phi(\lambda, q, \nu\sqrt{1-q}) &= \frac{\lambda^2}{2\pi(1-q)} \sum_{m=0}^{+\infty} \Delta x_m e^{-\lambda^2 x_m^2/2} x_m \sin(x_m \lambda \nu \sqrt{q}) \\ &\quad + \frac{\lambda \nu}{2\pi\sqrt{q}} \sum_{m=0}^{+\infty} \Delta x_m e^{-\lambda^2 x_m^2/2} \cos(x_m \lambda \nu \sqrt{q}), \end{aligned} \quad (95)$$

$$\mathcal{X}(\lambda, q, \nu\sqrt{1-q}) = \frac{1}{q\sqrt{1-q}} \sum_{m=0}^{+\infty} \Delta x_m e^{-\lambda^2 x_m^2/2} \cos(x_m \lambda \nu \sqrt{q}). \quad (96)$$

Notice that the discrete points x_m are in the middle of the intervals

$$\left[2m\sqrt{1-q}, 2(m+1)\sqrt{1-q} \right]. \quad (97)$$

It then follows that the series in (94)–(96) are Riemann sums such that, when $q \rightarrow 1^-$,

$$\Psi(\lambda, q, \nu\sqrt{1-q}) \simeq \frac{1}{2} + \frac{1}{\pi} \int_0^{+\infty} dy e^{-y^2/2} \frac{\sin(y\nu)}{y} = \frac{1}{2} \left(1 + \operatorname{erf}(\nu/\sqrt{2}) \right), \quad (98)$$

$$\begin{aligned}\Phi(\lambda, q, \nu\sqrt{1-q}) &\simeq \frac{1}{2\pi(1-q)} \int_0^{+\infty} dy e^{-y^2/2} y \sin(y\nu) + \frac{\nu}{2\pi} \int_0^{+\infty} dy e^{-y^2/2} \cos(y\nu) \\ &\simeq \frac{\nu}{2\sqrt{2\pi}(1-q)} e^{-\nu^2/2},\end{aligned}\quad (99)$$

$$\mathcal{X}(\lambda, q, \nu\sqrt{1-q}) \simeq \frac{1}{\lambda\sqrt{1-q}} \int_0^{+\infty} dy e^{-y^2/2} \cos(y\nu) = \frac{1}{\lambda} \sqrt{\frac{\pi}{2(1-q)}} e^{-\nu^2/2}, \quad (100)$$

where, in passing from the first line to the second one in equation (99), we have only retained the dominant diverging contribution when $q \rightarrow 1^-$ and, in (98), we have introduced the error function

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z dt e^{-t^2}. \quad (101)$$

Then, as $q \rightarrow 1^-$, the dominant behavior of $J(\lambda, q)$ in (93) is determined by (98) and (99), the latter diverging faster than (100), as $\frac{1}{1-q}$ against $\frac{1}{\sqrt{1-q}}$. Therefore,

$$\begin{aligned}J(\lambda, q) &\simeq (1-q)^{3/2} \sum_{n=0}^{+\infty} \int_0^{2\pi/(\lambda\sqrt{1-q})} \frac{d\nu}{2\pi} \nu \exp\left(-\frac{\nu^2}{2} - \frac{1}{2}\left(\nu\sqrt{1-q} + \frac{2\pi n}{\lambda}\right)^2\right) \\ &\quad \times \left(\frac{1}{1+\operatorname{erf}(\nu/\sqrt{2})} - \frac{1}{1-\operatorname{erf}(\nu/\sqrt{2})}\right).\end{aligned}\quad (102)$$

Going back to the integration variable $\omega = \nu\sqrt{1-q}$, one gets

$$\begin{aligned}J(\lambda, q) &\simeq (1-q)^{1/2} \sum_{n=0}^{+\infty} \int_0^{2\pi/\lambda} \frac{d\omega}{2\pi} \omega \exp\left(-\frac{\omega^2}{2(1-q)} - \frac{1}{2}\left(\omega + \frac{2\pi n}{\lambda}\right)^2\right) \\ &\quad \times \left(\frac{1}{1+\operatorname{erf}(\omega/\sqrt{2(1-q)})} - \frac{1}{1-\operatorname{erf}(\omega/\sqrt{2(1-q)})}\right).\end{aligned}\quad (103)$$

From the asymptotic behavior

$$\operatorname{erf}(z) \simeq 1 - \frac{1}{\sqrt{\pi}z} e^{-z^2} \quad \text{when} \quad z \rightarrow +\infty, \quad (104)$$

one obtains the expression (35); indeed,

$$\begin{aligned}J(\lambda, q) &\simeq (1-q)^{1/2} \sum_{n=0}^{+\infty} \int_0^{2\pi/\lambda} \frac{d\omega}{2\pi} \omega \exp\left(-\frac{\omega^2}{2(1-q)} - \frac{1}{2}\left(\omega + \frac{2\pi n}{\lambda}\right)^2\right) \\ &\quad \times \left(\frac{1}{2} - \sqrt{\frac{\pi}{2(1-q)}} \omega \exp\left(\frac{\omega^2}{2(1-q)}\right)\right) \\ &\simeq - \sum_{n=0}^{+\infty} \int_0^{2\pi/\lambda} \frac{d\omega}{2\sqrt{2\pi}} \omega^2 \exp\left(-\frac{1}{2}\left(\omega + \frac{2\pi n}{\lambda}\right)^2\right).\end{aligned}\quad (105)$$

Appendix C. Uniform convergence of the series in (35)

Setting $\frac{1}{\lambda} := s$, consider the sequence

$$\left\{f_k(s) := \int_0^{2\pi s} \frac{d\omega}{\sqrt{2\pi}} \omega^2 e^{-\frac{1}{2}(\omega+2\pi sk)^2}\right\}_k. \quad (106)$$

Taking the derivative with respect to s we obtain

$$\frac{d}{ds} f_k(s) = \sqrt{2\pi} (2\pi s)^2 e^{-\frac{1}{2}(2\pi s)^2(1+k)^2} - \int_0^{2\pi s} \frac{d\omega}{\sqrt{2\pi}} 2\pi k (\omega + 2\pi ks) \omega^2 e^{-\frac{1}{2}(\omega+2\pi sk)^2}. \quad (107)$$

Then, observing that

$$e^{-x^2} \leq \frac{1}{1+x^2}$$

and that the second integral is always positive,

$$\frac{d}{ds} f_k(s) \leq \frac{\sqrt{2\pi} (2\pi s)^2}{1 + 2\pi^2 s^2 (k+1)^2} \leq \frac{2\sqrt{2\pi}}{(1+k)^2}. \quad (108)$$

On the other hand, we know that

$$\frac{4\pi}{\sqrt{2\pi}} \sum_{k=0}^{\infty} \frac{1}{(1+k)^2} = \frac{4\pi}{\sqrt{2\pi}} \frac{\pi^2}{6} < +\infty, \quad (109)$$

therefore, using the Weierstrass M-test (see e.g. [38]), we can conclude that the sequence $\{\frac{d}{ds} f_k(s)\}_k$ is uniformly convergent in any open interval of the kind $(0, S)$, with $S < \infty$. It is also clear that the sequence $\{f_k(s)\}_k$ converges in $s=0$. Then, as a consequence of theorem 1.23 in [38], we have that $\{f_k(s)\}_k$ is uniformly convergent in any compact interval $[0, S]$.

Appendix D. Analytical approximation of $\Phi(\lambda, q, \omega)$

In the following, we provide an approximated expression of the function $\Phi(\lambda, q, \omega)$ introduced in (32) as a linear combination of special functions. As a byproduct, we will also derive another proof of the fact that $\lim_{\lambda \rightarrow +\infty} \alpha(\lambda) = +\infty$. The function $\Phi(\lambda, q, \omega) := \partial_q \Psi(\lambda, q, \omega)$ is given by

$$\begin{aligned} \partial_q \left(\sum_{k \in \mathbb{Z}} \int_{\epsilon_1(\lambda, q, \omega, k)}^{\epsilon_2(\lambda, q, \omega, k)} \frac{dz}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \right) &= \sum_{k \in \mathbb{Z}} \exp \left(-\frac{(\pi + 2k\pi - \lambda\sqrt{q}\omega)^2}{2\lambda^2(1-q)} \right) \left(\frac{-\omega\lambda + \sqrt{q}\pi(1+2k)}{2\lambda\sqrt{q}(1-q)^{3/2}} \right) \\ &\quad - \exp \left(-\frac{(2k\pi - \lambda\sqrt{q}\omega)^2}{2\lambda^2(1-q)} \right) \left(\frac{-\omega\lambda + 2k\pi\sqrt{q}}{2\lambda\sqrt{q}(1-q)^{3/2}} \right), \end{aligned} \quad (110)$$

where we have used the explicit expression of $\epsilon_1(\lambda, q, \omega, k)$ and $\epsilon_2(\lambda, q, \omega, k)$,

$$\epsilon_1(\lambda, q, \omega, k) := \frac{-\lambda\sqrt{q}\omega + 2k\pi}{\lambda\sqrt{1-q}}, \quad \epsilon_2(\lambda, q, \omega, k) := \frac{-\lambda\sqrt{q}\omega + \pi + 2k\pi}{\lambda\sqrt{1-q}}. \quad (111)$$

To construct the desired approximation of $\Phi(\lambda, q, \omega)$, we replace the Gaussian terms appearing in (110) with Lorentzian functions, namely, we set

$$\exp \left(-\frac{(\pi + 2k\pi - \lambda\sqrt{q}\omega)^2}{2\lambda^2(1-q)} \right) \simeq \frac{1}{1 + \frac{(\pi + 2k\pi - \lambda\sqrt{q}\omega)^2}{2\lambda^2(1-q)}} = \frac{2\lambda^2(1-q)}{2\lambda^2(1-q) + (\pi + 2k\pi - \lambda\sqrt{q}\omega)^2}, \quad (112)$$

and

$$\exp \left(-\frac{(2k\pi - \lambda\sqrt{q}\omega)^2}{2\lambda^2(1-q)} \right) \simeq \frac{1}{1 + \frac{(2k\pi - \lambda\sqrt{q}\omega)^2}{2\lambda^2(1-q)}} = \frac{2\lambda^2(1-q)}{2\lambda^2(1-q) + (2k\pi - \lambda\sqrt{q}\omega)^2}. \quad (113)$$

Note that, these approximation are the more accurate, the closer is q to 1 from below, or the greater is λ . Inserting (112) and (113) into (110) yields a new function $\tilde{\Phi}(\lambda, q, \omega)$:

$$\begin{aligned} \tilde{\Phi}(\lambda, q, \omega) &= \sum_{k \in \mathbb{Z}} \frac{\lambda}{\sqrt{q}\sqrt{1-q}} \left(\frac{-\omega\lambda + \sqrt{q}\pi(1+2k)}{2\lambda^2(1-q) + (\pi + 2k\pi - \lambda\sqrt{q}\omega)^2} \right. \\ &\quad \left. - \frac{-\omega\lambda + 2k\pi\sqrt{q}}{2\lambda^2(1-q) + (2k\pi - \lambda\sqrt{q}\omega)^2} \right). \end{aligned} \quad (114)$$

We now show that the series in (114) can be exactly summed, and provides an approximation of the function $\Phi(\lambda, q, \omega)$. Indeed, consider first the term

$$\begin{aligned}
& \frac{\lambda}{\sqrt{q}\sqrt{1-q}} \sum_{k \in \mathbb{Z}} \frac{-\omega\lambda + \sqrt{q}\pi(1+2k)}{2\lambda^2(1-q) + (\pi + 2k\pi - \lambda\sqrt{q}\omega)^2} \\
&= \frac{\lambda}{\sqrt{q}\sqrt{1-q}} \sum_{k \in \mathbb{Z}} \frac{-\omega\lambda + \sqrt{q}\pi(1+2k)}{\left(\pi + 2k\pi - \lambda\sqrt{q}\omega + \sqrt{2\lambda^2(q-1)}\right)\left(\pi + 2k\pi - \lambda\sqrt{q}\omega - \sqrt{2\lambda^2(q-1)}\right)} \\
&= \frac{2\lambda\sqrt{q} - \omega\sqrt{2\lambda^2(q-1)}}{8\pi\sqrt{q}\sqrt{1-q}} \sum_{k \in \mathbb{Z}} \frac{1}{k + \left(\frac{1}{2} + \frac{\sqrt{2\lambda^2(q-1)} - \lambda\sqrt{q}\omega}{2\pi}\right)} \\
&\quad + \frac{\omega\sqrt{2\lambda^2(q-1)} + 2\sqrt{q}\lambda}{8\pi\sqrt{q}\sqrt{1-q}} \sum_{k \in \mathbb{Z}} \frac{1}{k + \left(\frac{1}{2} - \frac{\lambda\sqrt{q}\omega + \sqrt{2\lambda^2(q-1)}}{2\pi}\right)}. \tag{115}
\end{aligned}$$

Let us consider separately the contribution to the series for $k \in [-\infty, -1]$ and $k \in [0, \infty]$. For the negative values of k , we have:

$$\begin{aligned}
& \frac{2\lambda\sqrt{q} - \omega\sqrt{2\lambda^2(q-1)}}{8\pi\sqrt{q}\sqrt{1-q}} \sum_{k=-\infty}^{-1} \frac{1}{k + \left(\frac{1}{2} + \frac{\sqrt{2\lambda^2(q-1)} - \lambda\sqrt{q}\omega}{2\pi}\right)} \\
&+ \frac{\omega\sqrt{2\lambda^2(q-1)} + 2\sqrt{q}\lambda}{8\pi\sqrt{q}\sqrt{1-q}} \sum_{k=-\infty}^{-1} \frac{1}{k + \left(\frac{1}{2} - \frac{\lambda\sqrt{q}\omega + \sqrt{2\lambda^2(q-1)}}{2\pi}\right)} \\
&= -\frac{2\lambda\sqrt{q} - \omega\sqrt{2\lambda^2(q-1)}}{8\pi\sqrt{q}\sqrt{1-q}} \sum_{k=1}^{+\infty} \frac{1}{k - \left(\frac{1}{2} + \frac{\sqrt{2\lambda^2(q-1)} - \lambda\sqrt{q}\omega}{2\pi}\right)} \\
&\quad - \frac{\omega\sqrt{2\lambda^2(q-1)} + 2\sqrt{q}\lambda}{8\pi\sqrt{q}\sqrt{1-q}} \sum_{k=1}^{+\infty} \frac{1}{k - \left(\frac{1}{2} - \frac{\lambda\sqrt{q}\omega + \sqrt{2\lambda^2(q-1)}}{2\pi}\right)}, \tag{116}
\end{aligned}$$

where we have changed the sign of k to $-k$. It is worth observing that the series in (116) is divergent, as one expects upon considering that the series (115) will converge only considering the values of $k \in [-\infty, +\infty]$. Indeed, sending k into $k+1$, and adding and subtracting the factor $\frac{1}{k+1}$ in (116), we get

$$\begin{aligned}
& -\frac{2\lambda\sqrt{q} - \omega\sqrt{2\lambda^2(q-1)}}{8\pi\sqrt{q}\sqrt{1-q}} \sum_{k=0}^{+\infty} \frac{1}{k + \left(\frac{1}{2} - \frac{\sqrt{2\lambda^2(q-1)} - \lambda\sqrt{q}\omega}{2\pi}\right)} \\
&\quad - \frac{\omega\sqrt{2\lambda^2(q-1)} + 2\sqrt{q}\lambda}{8\pi\sqrt{q}\sqrt{1-q}} \sum_{k=0}^{+\infty} \frac{1}{k + \left(\frac{1}{2} + \frac{\lambda\sqrt{q}\omega + \sqrt{2\lambda^2(q-1)}}{2\pi}\right)} \\
&= \frac{4\sqrt{q}\lambda}{8\pi\sqrt{q}\sqrt{1-q}} \gamma - \frac{4\sqrt{q}\lambda}{8\pi\sqrt{q}\sqrt{1-q}} \sum_{k=1}^{+\infty} \frac{1}{k+1} \\
&\quad + \frac{2\sqrt{q}\lambda - \omega\sqrt{2\lambda^2(q-1)}}{8\pi\sqrt{q}\sqrt{1-q}} \psi^{(0)}\left(\frac{1}{2} - \frac{\sqrt{2\lambda^2(q-1)} - \lambda\sqrt{q}\omega}{2\pi}\right) \\
&\quad + \frac{\omega\sqrt{2\lambda^2(q-1)} + 2\sqrt{q}\lambda}{8\pi\sqrt{q}\sqrt{1-q}} \psi^{(0)}\left(\frac{1}{2} + \frac{\lambda\sqrt{q}\omega + \sqrt{2\lambda^2(q-1)}}{2\pi}\right), \tag{117}
\end{aligned}$$

where γ is the so-called *Euler–Mascheroni* constant, while $\psi^{(0)}(z)$ is the *digamma function* [39, 40]—namely, the function defined as

$$\psi^{(0)}(z) := \frac{\Gamma'(z)}{\Gamma(z)} = -\gamma + \sum_{k=0}^{+\infty} \frac{1}{k+1} - \frac{1}{z+k}, \quad \operatorname{Re} z > 0. \quad (118)$$

As anticipated above, a divergent term appears in (117). Nevertheless, considering the contribution to the series (115) for $k \in [0, +\infty]$ yields

$$\begin{aligned} & \frac{2\lambda\sqrt{q} - \omega\sqrt{2\lambda^2(q-1)}}{8\pi\sqrt{q}\sqrt{1-q}} \sum_{k=0}^{+\infty} \frac{1}{k + \left(\frac{1}{2} + \frac{\sqrt{2\lambda^2(q-1)} - \lambda\sqrt{q}\omega}{2\pi}\right)} \\ & + \frac{\omega\sqrt{2\lambda^2(q-1)} + 2\sqrt{q}\lambda}{8\pi\sqrt{q}\sqrt{1-q}} \sum_{k=0}^{+\infty} \frac{1}{k + \left(\frac{1}{2} - \frac{\lambda\sqrt{q}\omega + \sqrt{2\lambda^2(q-1)}}{2\pi}\right)} \\ & = -\frac{4\lambda\sqrt{q}}{8\pi\sqrt{q}\sqrt{1-q}} \gamma + \frac{4\lambda\sqrt{q}}{8\pi\sqrt{q}\sqrt{1-q}} \sum_{k=0}^{\infty} \frac{1}{k+1} \\ & - \frac{2\lambda\sqrt{q} - \omega\sqrt{2\lambda^2(q-1)}}{8\pi\sqrt{q}\sqrt{1-q}} \psi^{(0)}\left(\frac{1}{2} + \frac{-\lambda\sqrt{q}\omega + \sqrt{2\lambda^2(q-1)}}{2\pi}\right) \\ & - \frac{\omega\sqrt{2\lambda^2(q-1)} + 2\sqrt{q}\lambda}{8\pi\sqrt{q}\sqrt{1-q}} \psi^{(0)}\left(\frac{1}{2} - \frac{\lambda\sqrt{q}\omega + \sqrt{2\lambda^2(q-1)}}{2\pi}\right). \end{aligned} \quad (119)$$

Combining equations (117) and (119), we obtain the final expression

$$\begin{aligned} & \frac{\lambda}{\sqrt{q}\sqrt{1-q}} \sum_{k \in \mathbb{Z}} \frac{-\omega\lambda + \sqrt{q}\pi(1+2k)}{2\lambda^2(1-q) + (\pi + 2k\pi - \lambda\sqrt{q}\omega)^2} \\ & = C_1(\lambda, q, \omega) \left(\psi^{(0)}\left(\frac{1}{2} + Z(\lambda, q, \omega)\right) - \psi^{(0)}\left(\frac{1}{2} - Z(\lambda, q, \omega)\right) \right) \\ & + C_2(\lambda, q, \omega) \left(\psi^{(0)}\left(\frac{1}{2} + W(\lambda, q, \omega)\right) - \psi^{(0)}\left(\frac{1}{2} - W(\lambda, q, \omega)\right) \right), \end{aligned} \quad (120)$$

where we set

$$\begin{aligned} Z(\lambda, q, \omega) &= \frac{\lambda\sqrt{q}\omega - \sqrt{2\lambda^2(q-1)}}{2\pi}, & W(\lambda, q, \omega) &= \frac{\lambda\sqrt{q}\omega + \sqrt{2\lambda^2(q-1)}}{2\pi} \\ C_1(\lambda, q, \omega) &= \frac{2\sqrt{q}\lambda - \omega\sqrt{2\lambda^2(q-1)}}{8\pi\sqrt{q}\sqrt{1-q}}, & C_2(\lambda, q, \omega) &= \frac{\omega\sqrt{2\lambda^2(q-1)} + 2\sqrt{q}\lambda}{8\pi\sqrt{q}\sqrt{1-q}}, \end{aligned} \quad (121)$$

and where, as expected, the divergent contributions are no longer present. Analogously, we can sum the second term in the series (114) thus arriving at the following final expression of the function $\tilde{\Phi}(\lambda, q, \omega)$:

$$\begin{aligned} \tilde{\Phi}(\lambda, q, \omega) &= C_1(\lambda, q, \omega) \left(\psi^{(0)}\left(\frac{1}{2} + Z(\lambda, q, \omega)\right) - \psi^{(0)}(1 + Z(\lambda, q, \omega)) + \psi^{(0)}(-Z(\lambda, q, \omega)) \right. \\ & \quad \left. - \psi^{(0)}\left(\frac{1}{2} - Z(\lambda, q, \omega)\right) \right) + C_2(\lambda, q, \omega) \left(\psi^{(0)}\left(\frac{1}{2} + W(\lambda, q, \omega)\right) - \psi^{(0)}\left(\frac{1}{2} - W(\lambda, q, \omega)\right) \right. \\ & \quad \left. + \psi^{(0)}(-W(\lambda, q, \omega)) - \psi^{(0)}(1 + W(\lambda, q, \omega)) \right). \end{aligned} \quad (122)$$

A close inspection of the function $\tilde{\Phi}(\lambda, q, \omega)$ shows that its modulus, $|\tilde{\Phi}(\lambda, q, \omega)|$, actually provides an upper bound to $\Phi(\lambda, q, \omega)$, which is consistent with the approximation of Gaussian functions by Lorentzian ones considered in (112) and (113).

Using the well known asymptotic expansion of the digamma function [40]:

$$\psi^{(0)}(z) \sim \ln z - \frac{1}{2z}, \quad |z| \rightarrow +\infty, \quad (123)$$

we can also derive the asymptotic expansion of the function $\tilde{\Phi}(\lambda, q, \omega)$, for $\lambda \rightarrow \infty$, i.e.

$$\tilde{\Phi}(\lambda, q, \omega) \stackrel{\lambda \rightarrow \infty}{\simeq} C_1(\lambda, q, \omega) \left(\ln \left(\frac{\pi - \sqrt{2\lambda^2(q-1)} + \lambda\sqrt{q}\omega}{2\pi + \lambda\sqrt{q}\omega - \sqrt{2\lambda^2(q-1)}} \right) \right. \quad (124)$$

$$\left. - \ln \left(\frac{\pi - \lambda\sqrt{q}\omega + \sqrt{2\lambda^2(q-1)}}{-\lambda\sqrt{q}\omega + \sqrt{2\lambda^2(q-1)}} \right) \right) + C_2(\lambda, q, \omega) \left(\ln \left(\frac{\pi + \lambda\sqrt{q}\omega + \sqrt{2\lambda^2(q-1)}}{\pi - \lambda\sqrt{q}\omega - \sqrt{2\lambda^2(q-1)}} \right) - \ln \left(\frac{2\pi + \lambda\sqrt{q}\omega + \sqrt{2\lambda^2(q-1)}}{-\lambda\sqrt{q}\omega - \sqrt{2\lambda^2(q-1)}} \right) \right), \quad (125)$$

from which, it is not difficult to show that $\lim_{\lambda \rightarrow \infty} \tilde{\Phi}(\lambda, q, \omega) = 0$. On the other hand, taking into account that $\lim_{\lambda \rightarrow \infty} \Psi(\lambda, q, \omega) = 1/2$ (see (82)), it follows that $\lim_{\lambda \rightarrow \infty} \tilde{\Phi}(\lambda, q, \omega)/\Psi(\lambda, q, \omega) = 0$, from which we can argue, once again, the limit $\lim_{\lambda \rightarrow \infty} \alpha(\lambda) = +\infty$.

References

- [1] Biamonte J, Wittek P, Pancotti N, Rebentrost P, Wiebe N and Lloyd S 2017 Quantum machine learning *Nature* **549** 195–202
- [2] Engel A and van den Broeck C 2000 *Statistical Mechanics of Learning* (Cambridge University Press)
- [3] Nishimori H 2001 *Statistical Physics of Spin Glasses and Information Processing: An Introduction* (Oxford University Press)
- [4] Gardner E 1987 Maximum storage capacity of neural networks *Europhys. Lett.* **4** 481–5
- [5] Gardner E 1988 The space of interactions in neural network models *J. Phys. A* **21** 257–70
- [6] Gardner E and Derrida B 1988 Optimal storage properties of neural network models *J. Phys. A* **21** 271–84
- [7] Benatti F, Gramegna G and Mancini S 2022 Pattern capacity of a single quantum perceptron *J. Phys. A: Math. Theor.* **55** 155301
- [8] Benatti F, Gramegna G, Mancini S and Nwemadji G 2024 On the capacity of a quantum perceptron for storing biased patterns *J. Phys. A: Math. Theor.* **57** 025301
- [9] Gratsea K, Kasper V and Lewenstein M 2024 Storage properties of a quantum perceptron *Phys. Rev. E* **110** 024127
- [10] Urushibata M and Ohzeki M 2025 Storage capacity evaluation of quantum perceptron by replica method *J. Phys. Soc. Japan* **94** 054002
- [11] Torrontegui E and García-Ripoll J J 2019 Unitary quantum perceptron as efficient universal approximator *Europhys. Lett.* **125** 30004
- [12] Rosenblatt F 1957 The perceptron: a perceiving and recognizing automaton *Technical Report* (Inc. Report No. 85-460-1) Cornell Aeronautical laboratory
- [13] Rosenblatt F 1958 The perceptron: a probabilistic model for information storage and organization in the brain *Psych. Rew.* **65** 386–408
- [14] McCulloch W and Pitts W 1943 A logical calculus of ideas immanent in nervous activity *Bull. Math. Biophys.* **5** 115–33
- [15] Goodfellow I, Bengio Y and Courville A 2016 *Deep Learning* (The MIT Press)
- [16] Hertz J and Palmer R G 2018 *Introduction to the Theory of Neural Computation* (CRC Press)
- [17] Mézard M, Parisi G and Virasoro M A 1987 *Spin Glass Theory and Beyond: An Introduction to the Replica Method and its Applications* (World Scientific Publishing Company)
- [18] Malatesta E 2023 High-dimensional manifold of solutions in neural networks: insights from statistical physics (arXiv:2309.09240)
- [19] Talagrand M 2011 *Mean Field Models for Spin Glasses. Volume I: Basic Examples* (Springer)
- [20] Watkin T, Rau A and Biehl M 1993 The statistical mechanics of learning a rule *Rev. Mod. Phys.* **65** 499
- [21] Cover T M 1965 Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition *IEEE Trans. Electron. Comput.* **14** 326–34
- [22] Benatti F, Mancini S and Mangini S 2019 Continuous variable quantum perceptron *Int. J. Quantum Inf.* **17** 1941009
- [23] Tóth G, Lent C S, Tougaw P D, Brazhnik Y, Weng W, Porod W, Liu R-W and Huang Y-F 1996 Quantum cellular neural networks *Superlattices Microstruct.* **20** 473–8
- [24] McClean J R, Romero J, Babbush R and Aspuru-Guzik A 2016 The theory of variational hybrid quantum-classical algorithms *New J. Phys.* **18** 023023
- [25] Romero J and Aspuru-Guzik A 2021 Variational quantum generators: generative adversarial quantum machine learning for continuous distributions *Adv. Quantum Technol.* **4** 2000003
- [26] Mitarai K, Negoro M, Kitagawa M and Fujii K 2018 Quantum circuit learning *Phys. Rev. A* **98** 032309
- [27] McClean J R, Boixo S, Smelyanskiy V N, Babbush R and Neven H 2018 Barren plateaus in quantum neural network training landscapes *Nat. Commun.* **9** 4812
- [28] Benedetti M, Lloyd E, Sack S and Fiorentini M 2019 Parameterized quantum circuits as machine learning models *Quantum Sci. Technol.* **4** 043001
- [29] Gupta S and Zia R K P 2001 Quantum neural networks *J. Comput. Syst. Sci.* **63** 355–83
- [30] Andrecut M and Ali M K 2002 A quantum neural network model *Int. J. Mod. Phys. C* **13** 75–88
- [31] Zhou R, Wang H, Wu Q and Shi Y 2012 Quantum associative neural network with nonlinear search algorithm *Int. J. Theor. Phys.* **51** 705–23
- [32] da Silva A J, Luderger T S and de Oliveira W R 2016 Quantum perceptron over a field and neural network architecture selection in a quantum computer *Neural Netw.* **76** 55–64
- [33] Wan K H, Dahlsten O, Kristjánsson H, Gardner R and Kim M S 2017 Quantum generalisation of feedforward neural networks *npj Quantum Inf.* **3** 36
- [34] Schuld M, Sweke R and Meyer J J 2021 Effect of data encoding on the expressive power of variational quantum-machine-learning models *Phys. Rev. A* **103** 032430
- [35] Cao Y, Guerreschi G G, Giacomo G and Aspuru-Guzik A 2017 Quantum neuron: an elementary building block for machine learning on quantum computers (arXiv:1711.11240)

- [36] Sopena J M, Romero E and Alquézar R 1999 Neural networks with periodic and monotonic activation functions: a comparative study in classification problems *9th Int. Conf. Artificial Neural Networks: ICANN'99, IET*, 0-85296-721-7
- [37] Fyodorov Y 2010 Multifractality and freezing phenomena in random energy landscapes: an introduction *Physica A* **389** 4229
- [38] Knapp A W 2005 *Basic Real Analysis* (Birkhäuser)
- [39] Lebedev N N 1972 *Special Functions and Their Applications* (Dover Publications)
- [40] Olver F W J 2014 *Asymptotics and Special Functions* (Academic)